图书情报工作
第65卷 第9期 2021年5月

融入评论指标的中文图书综合评价体系适应性研究

朱世琴　邱悦　陈红英
华东理工大学科技信息研究所　上海 200237

摘　要：[目的/意义]为改进当前图书评价体系在反映图书内容质量方面的不足，将图书评论指标纳入评价体系，并应用不同学科图书分析评价体系的适应性，探索中文图书质量评价。[方法/过程]从传统计量指标、Altmetrics 指标和图书评论指标中筛选出 16 个指标，构建以 CRITIC 方法赋权的中文图书综合评价体系，同时选取经济学、计算机科学和物理学图书进行实证研究，分析图书评论指标在图书质量评价中的效用，并检验该评价体系在评价不同学科图书方面的适应性。[结果/结论]实证结果表明，中文图书综合评价体系对不同学科图书具有较好的适应性，评价结果可以对总被引、入藏馆量反映的图书影响力和质量进行补充；图书评论指标可以反映不同学科图书读者的阅读倾向，有助于筛选出具有较高可读性、科学性与知识性的学科图书，为阅读推广、书目推荐等图书质量评价活动提供参考。

关键词：图书评价　质量评价　图书评论　综合评价体系　适应性
分类号：G250
DOI：10.13266/j.issn.0252－3116.2021.09.003

1　引言

图书是知识密集型的学术成果，对于文化和知识的传承具有重要意义。随着图书出版数量的不断增长，如何科学而有效地评价图书，成为出版业工作人员与学者们亟需解决的问题。与期刊不同，图书由于载体形式、出版时间、内容长度等因素的制约，其评价工作的发展相对滞后。互联网的发展极大地丰富了图书的评价指标，但现有研究中图书评价体系的评价指标多基于图书的外部特征，如引文数、借阅量、社交媒体提及等，仅反映了图书的受欢迎程度，评价体系缺乏质性评价指标，难以揭示图书的质量。笔者构建的中文图书综合评价体系，尝试在传统计量指标、Altmetrics 指标的基础上引入图书评论指标，通过读者的阅读评论反馈图书内容质量，评价图书影响的同时实现对图书内容质量的揭示，以期降低单一类型指标评价的片面性，为中文图书质量评价提供一种全面、可行的方案。

2　图书评价研究现状

通过阅读、梳理国内外以“图书评价”“著作评价”和“专著评价”等为主题的相关文献，发现当前国内外学者在图书评价方面的成果，主要涉及评价指标的完善和评价体系的构建。

2.1　基于传统计量指标的图书评价研究

考虑到单纯基于期刊的引文索引可能不足以评估不同学科图书的引文影响，K. Kousha 等[1]对社会科学、艺术学和人文科学等学科书籍在线引用数据展开研究，发现上述学科图书的在线引用数据可以为图书的引文影响评估提供支持。入藏馆量[2]代表了持有指定书籍的图书馆的数量，可用于反馈图书馆员对图书的有用性评价。一项与引文指标的相关性研究发现，入藏馆量指标可以揭示与引文不同的影响，并在一定程度上弥补引文分析的缺陷[3]。此外，L. Cabezas-Clavijo 等[4]的研究表明借阅量与引用数之间没有显著相关性，这意味着图书的借阅量可以对引文的影响进行补充评价。

2.2　基于 Altmetrics 指标的图书评价研究

Altmetrics 的提出引起了国内外学者广泛关注，该指标的加入改善了传统计量指标存在的时滞性和马太效应，为图书评价提供了新的方向。一项针对 Altmetrics 指标在人文学科出版物中覆盖范围的统计发现，

作者简介：朱世琴（ORCID：0000－0001－5646－9287），副主任，副研究馆员，硕士生导师，E-mail：shqzhu@ecust.edu.cn；邱悦（ORCID：0000－0003－0455－0504），硕士研究生；陈红英（ORCID：0000－0002－4075－3774），硕士研究生。
收稿日期：2020－12－25　修回日期：2020－02－27　本文起止页码：23－31　本文责任编辑：徐健

Twitter 可以作为 Altmetrics 数据的可靠来源,为人文学科书籍提供社交媒体提及数据[5]。在 Altmetrics 指标的应用价值上,A. A. Zuccala 等[6]发现历史类书籍的 Goodreads 读者评分可以揭示图书的影响力。此外,纯科学、社会科学、工程和医学等不同学科图书间引用和下载量的显著相关性表明,下载量大的学科图书具有较高参考价值[7]。X. Nan 等[8]对 CBKCI 中 1 000 本高被引书籍的引文指标和豆瓣阅读网站的 Altmetrics 指标展开研究,研究表明 Altmetrics 指标和引文指标反映了图书影响力的不同方面,并在学术评价上形成互补。

2.3 基于书评的图书评价研究

从书评角度开展的图书评价研究中,K. Kousha 等[9]基于人文艺术、社会科学、科学与医学领域图书的研究表明,在线书评倾向于反映书籍在教育和文化方面的社会影响,而非学术影响。针对 Goodreads 图书指标的一项研究,同样说明图书评论可以反馈图书更广泛的影响[10]。国内学者基于图书评论内容的研究较少,章成志等[11]分析了利用单一平台和多个平台评论数据进行图书影响力评价的差异,发现整合电商、社交媒体等不同评论平台的读者评论数据,可以更加全面地评价图书的影响力。

2.4 图书评价体系研究

随着图书评价研究的深入,一些学者通过组合不同来源的评价指标,构建图书评价体系以综合评价图书的影响力。李雁翎等[12]提出了基于作者、出版社、图书馆、销售和网络舆情 5 个维度的图书评价体系及分析模型;刘利等[13]采用专家调查法获取图书质量评价指标和权重值,构建基于核心作者、核心出版社、作者所属研究机构、图书版次和书评利用等维度的外文学术图书质量评价模型;杨毓丽等[14]利用 Harmonic 算法计算合著者对图书的贡献,综合考量作者学术声望、图书出版流通情况与图书评介情况,构建外文图书质量评价体系,为单本外文图书评价提供参考;何峻等[15]运用层次分析法、专家调查法和模糊隶属度分析法,从图书内容质量和编辑质量两个层面建立中文图书综合评价体系,并依据评价结果确定各学科的核心图书和扩展图书,以供后续研究参考和借鉴;肖敖夏等[16]应用学术图书的 Bookmetrix 指标数据对学术图书的影响力展开评价,发现 Bookmetrix 指标数据对学术图书的影响力评价有补充价值,但也存在单纯使用频次统计导致学术成果的影响力和价值失真,以及学科差异较大等问题,建议后续研究纳入更多的指标数据开展更全面的分析。

2.5 图书评价研究不足

总之,现有图书评价研究多基于可直接获取的定量指标开展,如使用引文、借阅量、图书馆藏量、Altmetrics 数据等指标数据,仅有少数研究对内容数据进行细粒度挖掘[17],致使研究数据存在局限,少有研究能实现图书质量评价,特别是能反映图书质量的综合评价。笔者在图书计量指标的基础上增加评论指标,采用情感倾向分析和评论属性挖掘处理图书评论数据,通过图书评论包含的图书属性和情感倾向揭示图书质量,并使用不同学科图书数据开展实证分析,以期检验评价体系在不同学科图书评价中的适应性,为中文图书质量评价提供参考。

3 中文图书综合评价体系

3.1 评价指标选取规则

笔者在中文图书综合评价体系的构建过程中,主要依据以下 3 个规则选取评价指标:

(1)借鉴已有研究中的成熟指标。考虑到已有研究中的评价指标往往具有较高的研究价值,可以为评价体系中评价指标的选取提供参考,因此,评价指标的选择应基于现有研究成果,确保所选指标的科学性与合理性。

(2)评价指标应覆盖图书质量评价的各个方面。为了保证评价体系能够筛选出具有较高影响力与质量的图书,研究选取的评价指标应反映图书在不同层面、不同角度的价值。同时,各指标应相互补充、取长补短,以保障图书质量评价工作的全面开展。

(3)评价指标数据的可获取性。指标数据的可获得性是维护评价体系运行的关键,在指标选取的过程中应考虑到数据来源是否可靠、数据搜集是否方便,保障指标数据获取的完整性。

3.2 评价体系构建

基于上述规则,笔者从传统计量指标、Altmetrics 指标和图书评论指标中筛选出 16 个二级指标,并从数据来源角度将 16 个二级指标划分为图书利用、图书关注与图书评论 3 个维度,最终构建的中文图书综合评价体系见图 1。

在一级指标中,图书利用指标由传统计量指标组成,倾向于反映图书的学术影响力,包含已有研究中使用频次较高的总被引、入藏馆量与借阅次数指标,以及旨在反映出版时间较短图书引用情况的年均被引指标,以上指标是图书在学术利用、入藏和阅读过程中的使用数据,反映了图书在学术领域和图书馆界的认可

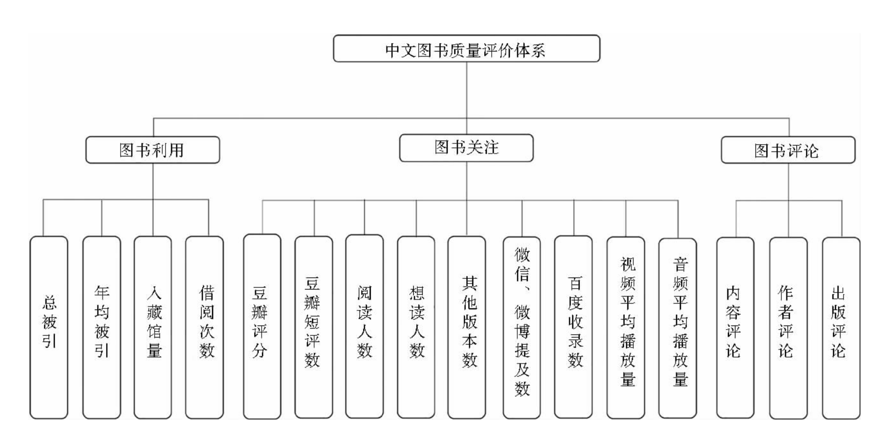


图1　中文图书综合评价体系

度,以及图书在读者间的受欢迎程度。为保障指标数据的可获得,选定数据来源为读秀网站和浙江图书馆。

图书关注指标能够反映图书的社会影响力,由豆瓣评分、豆瓣短评数、阅读人数、想读人数、其他版本数、微信和微博提及数、百度收录数、视频平均播放量和音频平均播放量等 Altmetrics 指标构成,是读者在社交网络环境下访问图书相关信息产生的行为数据,指标数据来自豆瓣读书、微信、微博、百度、bilibili 和喜马拉雅等网站。作为国内最大的阅读分享网站,豆瓣读书拥有较高的用户活跃度与知名度[18],由该网站提供的豆瓣评分、豆瓣短评数、阅读人数、想读人数和其他版本数可以呈现读者对图书的推荐度、图书在读者间传播的深度与广度以及图书在销售市场中的受欢迎程度;微信、微博提及数可以反映图书在大众读者间的流行趋势,弥补引用、借阅等指标数据的滞后性;百度收录数为图书的电子化版本在百度搜索引擎中被提及的次数,反映了电子版图书的影响力;视频和音频平均播放量为新增指标,反映在新型媒体阅读情境下,影视和有声书媒介对图书内容信息传播的影响,有助于更全面地衡量图书的影响力。

图书评论指标能够揭示图书的质量,笔者利用句末标点(“。”、“.”、“!”、“!”、“?”、“?”、“…”等)切分获得的图书评论,通过百度情感倾向分析 API 判断切分评论的情感倾向,并分别统计内容、作者和出版属性词在切分评论中的出现频次,当切分句中出现同一属性的不同属性词时,不进行重复计数(示例见表1),最终得到的指标数据为3类属性词在正、负向评论中出现频次的差值。评论属性词的设置参考已有研究[11],并根据现有评论中的高频词分布,将评论属性归为内容、作者和出版3类。内容属性包含部分、理论、感觉、内容、思维、观点等与图书内容相关的属性词,作者属性由作者、语言、文笔、文字等属性词组成,出版属性则涵盖了编辑、出版、翻译、译者、价格、包装、印刷、字体、封面、纸张、版本、插画等图书出版编辑相关属性词,旨在从图书的内容质量、作者文笔和出版质量3个方面反馈读者意见。

表1　评论属性统计举例

图书评论	情感倾向	内容属性	作者属性	出版属性
[作者]写的生动,[译者][翻译]的准确,的确是值得一读再读的好书	正	0	1	1
回顾了建国以来中国经济发展的起起伏伏,[内容]挺多,有些[观点]也挺新颖	正	1	0	0
中文版[翻译]的很一般,有的地方文绉绉的,硬伤也不少	负	0	0	1

3.3　CRITIC 法

笔者采用客观赋权方法中的 CRITIC 法计算指标权重,该方法将指标数据的变异程度和指标间的相关性纳入权重计算范畴,是一种优于熵权法和标准离差法的客观赋权方法,采用此方法得到的指标权重更符合客观实际,减少了主观因素对图书质量评价的干扰。应用 CRITIC 法计算指标权重时,首先依据指标间的相关系数计算冲突性,设 C_j 为第 j 个指标与其余指标的冲突性,则有:

$$C_j = \sum_{k=1}^{n}(1 - r_{jk}) \qquad 公式(1)$$

该指标的信息量为冲突性与标准差的乘积,具体如下:

$$I_j = \delta_j C_j \qquad 公式(2)$$

最终得到指标权重 W_j 为:

$$W_j = \frac{I_j}{\sum_{j=1}^{n} I_j} \qquad 公式(3)$$

由于本文的数据不符合正态分布,因此采用 Spearman 相关性分析获取各指标间的相关系数。

4 实证分析

4.1 数据来源

为探索不同学科图书应用中文图书综合评价体系的评价效果,并检验评价体系的适应性,笔者选择 3 个研究领域存在差异的学科图书进行实证研究,选取的学科包括经济学、计算机科学和物理学,分别来自社会科学、信息科学和自然科学领域。

笔者于 2020 年 2 月 1 日使用 Python 程序爬取豆瓣读书网站上标签为“经济学”“计算机科学”和“物理学”的图书,且 3 个学科中综合排名在前 20% 的图书,得到经济学图书 199 本、计算机科学图书 200 本和物理学图书 200 本。

截至 2020 年 2 月 10 日,通过 Python 程序爬取了 3 个学科图书的全部指标数据,剔除掉重复图书、指标数据缺失图书和合集图书等,共筛选得到经济学图书 140 本、计算机科学图书 137 本和物理学图书 136 本。

4.2 指标权重计算

通过 CRITIC 方法计算不同学科图书的二级指标权重,得到:

经济学 ω_1 = (0.032 5 0.032 7 0.078 5 0.059 1 0.115 8 0.044 2 0.044 9 0.059 8 0.043 9 0.065 1 0.043 8 0.054 6 0.053 9 0.084 3 0.078 9 0.108)

计算机科学 ω_2 = (0.041 4 0.033 2 0.096 7 0.085 1 0.104 7 0.032 0.034 2 0.044 9 0.111 7 0.059 4 0.041 1 0.051 2 0.044 7 0.079 1 0.062 7 0.0779)

物理学 ω_3 = (0.045 1 0.043 8 0.098 3 0.084 4 0.087 0.042 0.032 3 0.051 3 0.079 5 0.036 5 0.063 5 0.064 9 0.062 4 0.076 3 0.064 7 0.068)

4.3 评价结果分析

采用 min-max 标准化对样本图书的指标数据进行标准化处理,公式如下:

$$X'_{ij} = \frac{x_{ij} - \min_{1 \leqslant i \leqslant m} x_{ij}}{\max_{1 \leqslant i \leqslant m} x_{ij} - \min_{1 \leqslant i \leqslant m} x_{ij}} \quad i = 1, 2, \cdots, m \quad j = 1, 2, \cdots, n \qquad \text{公式 (4)}$$

其中,n 为评价指标个数,m 为评价对象个数。

4.3.1 综合得分分析

将标准化后的图书数据代入评价体系,得到 3 个学科图书的综合得分分布,如图 2 所示:

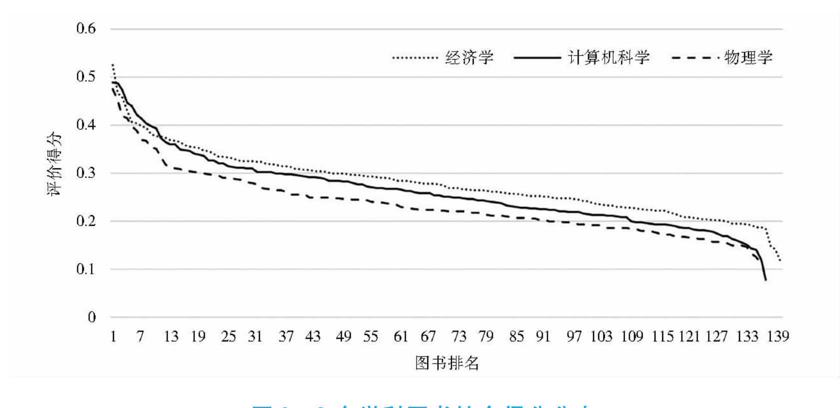


图 2 3 个学科图书综合得分分布

从图 2 可以看出,3 个学科图书的综合得分曲线均呈长尾分布。在评价得分为 0.3 至 0.5 的区间内聚集的图书约占 3 个学科整体书目的 20%,这时曲线的下降速率较快,不同学科图书均存在较大的得分差异;多数图书的评价得分在 0.1 至 0.3 的区间内,此时,经济学图书得分高于计算机科学图书与物理学图书,得分曲线的斜率缓于上一区间,图书间得分差距较小。总体而言,3 个学科图书的评价表现符合“少数图书获得较多关注,而多数图书表现平平”的市场规律。

从 3 个学科图书评价得分 Top10(见表 2)可以看出:①多数图书的图书关注得分较高,说明综合评价排名靠前的学科图书受到阅读网站、社交媒体和音频、视频媒体平台读者的关注较多。②图书主题与人们生产、生活相对接近,如经济学图书内容涵盖经济史、经济学概念、金融理财等,表明读者倾向于通过学习经济学基础知识、经济学史等掌握经济规律,并得到投资与理财方面的建议;计算机科学图书涉及计算机软硬件、数学思想、人物传记等,多数图书介绍了编程语言、算法思想、操作系统等实用知识与技术,反映读者阅读计算机科学书籍的主要目的在于获取技术知识;物理学书籍涉及天体物理、量子理论、物理学历史、基础物理等领域,说明读者更关注非常识的物理学知识,对以宇宙为代表的未知领域的兴趣是其阅读物理学图书的主要原因。

表2　3个学科图书评价(Top10)得分统计

学科	排名	书名	作者	评价得分	图书利用	图书关注	图书评论
经济学	1	激荡三十年	吴晓波	0.526	0.079	0.257	0.19
	2	富爸爸,穷爸爸	罗伯特・清崎	0.467	0.072	0.253	0.142
	3	国富论	亚当・斯密	0.459	0.064	0.257	0.138
	4	伟大的博弈	约翰・戈登	0.429	0.085	0.17	0.174
	5	投资中最简单的事	邱国鹭	0.407	0.06	0.135	0.212
	6	经济学原理	格里高利・曼昆	0.404	0.026	0.193	0.185
	7	金融的逻辑	陈志武	0.397	0.072	0.106	0.219
	8	历代经济变革得失	吴晓波	0.394	0.057	0.141	0.196
	9	货币战争	宋鸿兵	0.381	0.091	0.189	0.101
	10	变革中国	罗纳德・哈里	0.376	0.072	0.102	0.202
计算机科学	1	数学之美	吴军	0.49	0.151	0.228	0.111
	2	机器学习	周志华	0.487	0.23	0.123	0.134
	3	Java 编程思想	Bruce Eckel	0.474	0.182	0.191	0.101
	4	鸟哥的 Linux 私房菜	鸟哥	0.446	0.147	0.183	0.116
	5	黑客与画家	保罗・格雷厄姆	0.441	0.092	0.226	0.123
	6	高性能 MySQL	施瓦茨	0.423	0.137	0.149	0.137
	7	算法导论	Thomas H. Cormen	0.413	0.122	0.197	0.094
	8	数据结构与算法分析	Mark Allen Weiss	0.405	0.132	0.203	0.07
	9	线性代数及其应用	Lay D. C.	0.398	0.101	0.175	0.122
	10	Python 编程	埃里克・马瑟斯	0.394	0.126	0.142	0.126
物理学	1	时间简史	史蒂芬・霍金	0.476	0.066	0.302	0.108
	2	从一到无穷大	乔治・伽莫夫	0.452	0.141	0.186	0.125
	3	自然哲学之数学原理	牛顿	0.42	0.122	0.205	0.093
	4	相对论	爱因斯坦	0.414	0.172	0.133	0.109
	5	上帝掷骰子吗	曹天元	0.398	0.103	0.178	0.117
	6	天体运行论	哥白尼	0.389	0.158	0.145	0.086
	7	物理定律的本性	理查德・费曼	0.368	0.102	0.122	0.144
	8	黑洞与时间弯曲	基普.S.索恩	0.368	0.073	0.149	0.146
	9	超越时空	加来道雄	0.353	0.069	0.127	0.157
	10	七堂极简物理课	卡洛・罗韦利	0.352	0.091	0.14	0.121

4.3.2　图书评论指标分析

为探究图书评论指标在中文图书综合评价体系中的应用价值,依据3个学科图书的综合评价得分排名绘制图书评论得分曲线分布,如图3所示。可以看出,经济学图书的图书评论得分明显高于计算机科学和物理学图书,说明图书评论指标对经济学图书的影响较大,该学科图书的评论指标权重和短评数量位于3个学科前列。同时,排名前50%的学科图书其评论得分相对集中,位于其后的图书评论得分呈现随排名下降而递减的趋势,表明图书评论得分较高的图书其综合评价得分也相对较高,评价体系可以筛选出读者认可的学科图书。

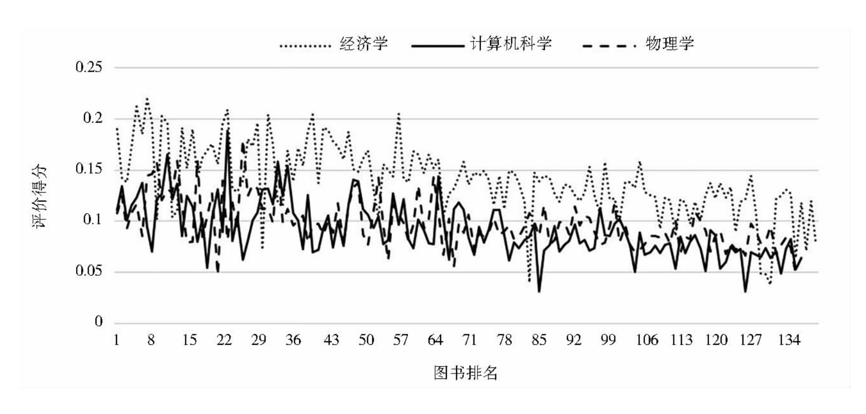


图3　3个学科图书评论指标得分

4.3.3 图书评论指标对图书排名影响分析

将 3 个学科指标数据代入无图书评论指标的评价体系中，统计图书评论指标增加前后两次评价排名变化情况，发现不同学科图书中排名发生变化的图书占全部图书 94% 以上，说明图书评论指标对不同学科图书均存在较大影响，列出排名差异 Top10 的图书，如表 3 所示：

表 3　图书评论指标增加前后 3 个学科图书排名差 Top10

经济学				计算机科学				物理学			
书名	增加前	增加后	排名差	书名	增加前	增加后	排名差	书名	增加前	增加后	排名差
经济学通识课	111	40	71	代码之髓	102	23	79	宇宙	88	26	62
战后日本经济史	125	57	68	STL 源码剖析	116	65	51	十问	120	64	56
斯坦福极简经济学	94	31	63	七周七语言	84	35	49	迷人的温度	87	33	54
城市的胜利	106	47	59	码农翻身	81	33	48	给忙碌者的天体物理学	108	61	47
棉花帝国	97	42	55	深度学习入门	83	48	35	一想到还有 95% 的问题留给人类，我就放心了	99	53	46
自私的皮球	98	43	55	穿越计算机的迷雾	89	56	33	物理学和哲学	89	49	40
通往奴役之路	10	83	-73	Python 深度学习	82	49	33	现实不似你所见	49	17	32
贫穷的本质	55	130	-75	计算机网络	66	104	-38	开启你的高维智慧	78	48	30
黑天鹅	49	129	-80	Java 核心技术・卷 I	78	126	-48	从混沌到有序	21	55	-34
随机漫步的傻瓜	37	131	-94	深度学习	33	85	-52	图解时间简史	31	68	-37

观察表 3 中不同学科图书的读者评论可以发现，排名下降的图书不存在学科差异，读者评论显示在翻译、内容逻辑和实用性方面有所欠缺。对排名上升的不同学科图书，经济学图书的多数读者认为其内容逻辑严密、具有实用性、阅读后读者有所启发；计算机科学图书多为入门类书籍，作者写作语言风趣并能够将理论与实践相结合，有助于读者学习计算机技能；物理学图书则为科普经典，兼具了科普与学术，内容有趣，能够引起读者思考。表 4 列出排名变化最大的 3 个学科图书评论：

表 4　3 个学科排名变化最大的图书评论示例

学科	排名	书名	评论
经济学	上升	经济学通识课	非常棒的经济学入门读本 这本书很详细介绍了经济的发展，通过一个个小故事介绍经济的一个个理论，让我们看到了经济学的方方面面 极好的一本经济学科普书，介绍经济学诞生与演变的重要观点，言语平实易懂
	下降	随机漫步的傻瓜	翻译确实不怎样，看原版应该更好 除了幸存者偏差理解更多一点外，其余一无所获 全书观点非常简单——随机性的重要性，但是全书的论述逻辑感觉是东一榔头西一棒槌，逻辑不够通畅
计算机科学	上升	代码之髓	虽然实际讨论的内容并不丰富，不过在一本薄薄的小册子里能够做到涵盖语言设计的核心要点，且行文流畅引人入胜，也是作者功底的体现 旁征博引的一本书，前面看得较仔细，后面速翻了一下通识书，非常适合初学编程者学习语言时配合阅读
	下降	深度学习	一星是给翻译的，感觉译者就没看懂，我是从第 14 章开始读的。前面的章节都看过英文版了，不夸张的说很多句子和机器翻译的没什么两样 这本书其实写的不好，不懂得人想通过这本书入门几乎不可能 翻译真的有问题，把行文逻辑都搞乱了
物理学	上升	宇宙	结合《宇宙》电视剧版和《宇宙时空之旅》食用更佳哦~这个版本是目前最良心的汉译本，译者本身就具有较强的专业知识，更具备对科普事业的热情 不止是科普著作，更给人以情感层面的震撼 翻译也不错
	下降	图解时间简史	目录编排混乱，语言就像网络上东拼西凑而成 总之这本书和霍金没有任何一点关系，纯粹胡编瞎扯，真不如《相对论史话》和《上帝掷骰子吗》 整体编排比较乱，有很多重复且前后矛盾的地方，脑子更乱了，原版还是得看

从表 4 中可以看出，经济学图书的读者主要关注经济学知识的获得，这类知识通常是以理论形式存在，因此多数读者对于经济学图书的逻辑框架、可读性、知识性具有较高需求；由于计算机科学图书多为技术性内容，读者希望通过阅读这类书籍获得技能相关知识，所以，内容偏基础、可读性较高、理论与实践兼备的计算机科学图书更为读者喜爱；物理学图书的内容涉及基础物理、天体物理等专业知识，读者阅读此类图书多

出于对物理知识的兴趣，这要求物理学图书具有较高的可读性和趣味性。

4.3.4 图书评论属性细分研究

为探究不同学科图书的读者对学科图书属性的评论倾向，笔者对内容评论和出版评论的属性词进一步划分，对作者评论属性不作区分，仅以“文笔”概述该属性的描述内容，划分维度如表5，并统计不同学科图书中二级属性词的分布情况，如图4所示：

表5 评论属性二级分类

一级属性	二级属性	属性词
内容评论	思想	思维、观点、逻辑、思考、思想
	事例	部分、理论、内容、概念、教材、书籍、本书、读物、方法、例子、知识、案例、道理、方式、原理、故事、章节、类型
	感受	感觉、理解、了解
作者评论	文笔	作者、语言、文笔、文字
出版评论	翻译	翻译、译者
	出版	编辑、出版、价格、包装、印刷、字体、封面、纸张、版本、插画

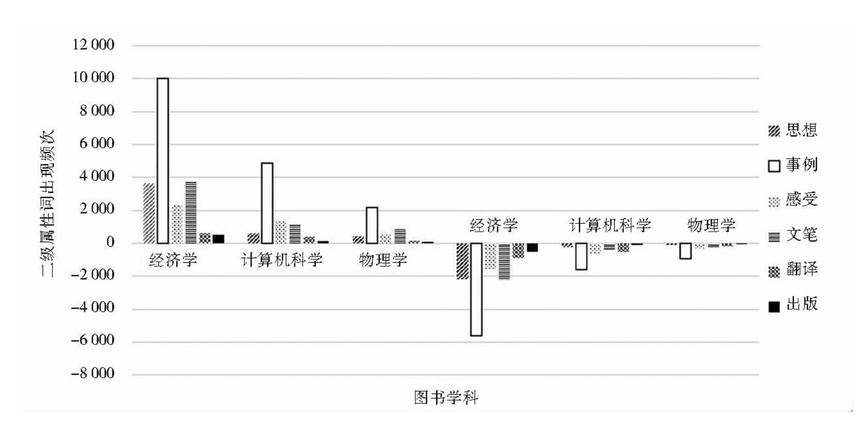


图4 图书评论二级属性分布

可以看出，不同学科图书的读者都对图书中包含的事例投入较多关注，表明读者希望通过学习书中的案例获得学科知识。除此之外，经济学、物理学图书的读者对作者文笔的评论较为积极，计算机科学图书的评论涉及读者主观感受较多，反映出读者在阅读社科类和科普类书籍时更关注阅读是否流畅，而在阅读技术类书籍时更在意对图书内容是否理解。比较不同学科图书的正、负向评论中二级属性出现频次发现，针对翻译属性的负面评价较多，且呈现出经济学、计算机科学、物理学图书依次递减的趋势，鉴于3个学科图书中外文译作的比例都在75%以上，且图书翻译水准对书籍的可读性和逻辑性影响较大，选择高水平的译著对于出版机构而言尤为重要。

基于上述分析，笔者认为图书评论指标可以揭示不同学科图书读者的阅读倾向，评价体系引入图书评论指标，有助于筛选出读者喜欢且具有较高的可读性、科学性及知识性的图书。

4.4 评价体系适应性分析

考虑到总被引指标和入藏馆量指标在图书影响力与质量揭示方面的权威性，笔者将3个学科图书评价结果分别与总被引和入藏馆量指标作相关性分析，以检验评价体系对于不同学科图书的影响力与质量的评价效果，判断评价体系在评价不同学科图书的适应性。

在进行相关性分析之前，首先利用Shapiro-Wilk检验判断3个学科图书评价结果与入藏馆量、总被引指标的分布情况，发现上述数据均不服从正态分布，因此笔者采用Spearman相关性检验分析指标间的相关性，进行相关性分析时的各学科图书数量为$N_{经济学}=140$，$N_{计算机科学}=137$，$N_{物理学}=136$，分析结果见表6。

相关性分析结果表明，3个学科图书的评价结果与入藏馆量和总被引指标均在0.01显著性水平呈正相关，相关程度中等偏弱，说明经济学、计算机科学、物理学图书的评价结果可以对入藏馆量指标反映的图书影响与价值，以及总被引指标反映的图书学术影响进行补充，对于图书书目推荐等图书质量评价活动具有一定的参考意义。

通过分析不同学科图书的评价结果与总被引、入藏馆量指标的相关性，笔者认为，中文图书综合评价体系可用于不同学科图书的质量评价，帮助读者与图书馆、阅读网站等图书收藏、阅读机构筛选出具有较高影响力与内容质量的图书。

表 6 3 个学科图书评价结果与入藏馆量、总被引指标的相关性分析结果

斯皮尔曼 Rho		评价结果		入藏馆量		总被引	
		相关系数	Sig.(双尾)	相关系数	Sig.(双尾)	相关系数	Sig.(双尾)
经济学	评价结果	1	.	.509 **	0.000	.491 **	0.000
	入藏馆量	.509 **	0.000	1	.	.652 **	0.000
	总被引	.491 **	0.000	.652 **	0.000	1	.
计算机科学	评价结果	1	.	.533 **	0.000	.437 **	0.000
	入藏馆量	.533 **	0.000	1	.	.458 **	0.000
	总被引	.437 **	0.000	.458 **	0.000	1	.
物理学	评价结果	1	.	.584 **	0.000	.462 **	0.000
	入藏馆量	.584 **	0.000	1	.	.395 **	0.000
	总被引	.462 **	0.000	.395 **	0.000	1	.

注:**. 在 0.01 级别(双尾),相关性显著

5 结语

笔者选取研究领域不同的学科图书,并分析学科图书在一级指标得分、图书评论指标效用方面的差异,以及评价结果与总被引、入藏馆量指标的相关性,展开评价体系对不同学科图书评价的适应性研究。根据实证结果,可以得出如下结论:

5.1 评价体系对不同学科图书具有较好的适应性

通过将不同学科图书的指标数据代入评价体系,并对评价结果与总被引、入藏馆量指标进行 Spearman 相关性分析,发现评价结果与总被引、入藏馆量指标均呈相关程度中等偏弱的显著性相关,说明评价体系可以反映不同学科图书的影响力和质量,并且增加的评论指标将读者对图书的情感和评价融入评价体系,揭示了图书在读者眼中的价值。本评价体系可应用于不同学科图书的评价,筛选出读者评价和学术质量、社交媒体影响力较高的图书,为阅读推广、书目推荐等图书质量评价活动提供参考。

5.2 图书评论指标可以反映不同学科图书读者的阅读倾向

对图书评论指标增加前后排名有显著性差异的图书评论分析发现,经济学图书偏重理论与经验的传播,其读者更关注图书的内在逻辑和知识性;计算机科学图书属于技术类图书,且涉及软件知识较多,这类书籍的读者希望在阅读中掌握一定的专业技能,所以偏重基础、可读性强、理论与实践相结合的书籍更受读者喜爱。上述两学科图书的读者更加青睐内容实用的图书,而定位为科普图书,且内容具有一定趣味性的物理学图书更易得到读者好评。

5.3 读者更多关注图书案例和翻译水平

统计图书评论的二级属性在正、负向评论中的出现频次发现,事例属性在不同学科图书评论中均保持较高的出现频次,说明读者对书中的具体事例较为关注,希望通过学习书中的案例获得学科知识。同时,针对翻译属性的负面评价较多,鉴于 3 个学科图书中外文译作比例在 75% 以上,且图书翻译水准对书籍的可读性和逻辑性影响较大,选择高水平的译著对于出版机构而言尤为重要。

由于技术条件和时间限制,笔者选择的传统计量指标、Altmetrics 指标和图书评论指标数据来源相对局限,其中图书评论数据仅采用了豆瓣读书网站的短评,后续研究中可以增加对引文内容、社交媒体提及内容的情感倾向分析和属性挖掘,以衡量不同来源的读者对图书的评价态度和关注属性,使图书质量评价更趋于完善。

参考文献:

[1] KOUSHA K, THELWALL M, REZAIE S. Assessing the citation impact of books: the role of Google books, Google Scholar, and Scopus[J]. Journal of the American Society for Information Science and Technology, 2011, 62(11):2147-2164.
[2] WHITE H D, BOELL S K, YU H, et al. Libcitations: a measure for comparative assessment of book publications in the humanities and social sciences[J]. Journal of the American Society for Information Science and Technology, 2009, 60(6):1083-1096.
[3] LINMANS A J M. Why with bibliometrics the Humanities does not need to be the weakest link[J]. Scientometrics, 2010, 83(2): 337-354.
[4] CABEZAS-CLAVIJO L, ROBINSON-GARCÍA N, TORRES-SALINAS D, et al. Most borrowed is most cited? library loan statistics as a proxy for monograph selection in citation indexes[EB/OL].

[2020－11－01]. https://arxiv. org/abs/1305. 1488.

[5] BJÖRN H. Using Altmetrics for assessing research impact in the humanities[J]. Scientometrics, 2014, 101(2):1419－1430.

[6] ZUCCALA A A, VERLEYSEN F T, CORNACCHIA R, et al. Altmetrics for the humanities: comparing goodreads reader ratings with citations to history books[J]. Aslib journal of information management, 2015, 67(3): 320－336.

[7] ERFANMANESH M, NOORHIDAWATI A, ABRIZAH A. What can Bookmetrix tell us about the impact of Springer Nature's books [J]. Scientometrics, 2019, 121(1):521－536.

[8] NAN X, LI M, SHI J. Using altmetrics for assessing impact of highly-cited books in Chinese Book Citation Index[J]. Scientometrics, 2020, 122(3):1651－1669.

[9] KOUSHA K, THELWALL M. Can Amazon. com reviews help to assess the wider impacts of books? [J]. Journal of the Association for Information Science and Technology, 2016, 67(3): 566－581.

[10] KOUSHA K, THELWALL M, ABDOLI M. Goodreads reviews to assess the wider impacts of books[J]. Journal of the Association for Information Science & Technology, 2017,68(8):2004－2016.

[11] 章成志,童甜甜,周清清. 整合不同评论平台的图书综合影响力评价研究[J]. 情报学报, 2018,37(9):861－873.

[12] 李雁翎,孙晓慧,陈玖冰. 五维图书评价体系及分析模型的建构[J]. 情报科学,2013, 31 (8):77－80,140.

[13] 刘利,袁曦临. 外文学术图书质量评价实证研究[J]. 图书情报工作,2011(21):93－97, 102.

[14] 杨毓丽,丁媛,张苏. 基于 ESI 单本外文图书质量评价体系实证研究——以计算机学科为例[J]. 图书馆杂志,2015(7):11－18.

[15] 何峻,蔡蓉华. 中文图书评价体系研究[J]. 大学图书馆学报,2016,34(3):51－58,15.

[16] 肖敖夏,杨思洛. Bookmetrix 指标在学术图书影响力评价中的应用[J]. 图书馆论坛, 2020,40(9):79－85.

[17] 周清清,章成志. 图书影响力评价研究进展与展望[J]. 大学图书馆学报,2020,38(3):24－33.

[18] 李明,李江,陈铭,等. 中文学术图书引文量与 Altmetrics 指标探索性分析及其启示[J]. 情报学报,2019,38(6):557－567.

作者贡献说明:

朱世琴:提出研究思路与方法,论文指导与修改;

邱悦:数据搜集、整理与分析,论文撰写;

陈红英:参与数据搜集与整理。

Research on the Adaptability of Chinese Book Comprehensive Evaluation System Incorporated with Review Indicators

Zhu Shiqin　Qiu Yue　Chen Hongying

Institute of Science & Technology Information, East China University of Science & Technology, Shanghai 200237

Abstract: [**Purpose/significance**] To improve the inadequacy of the current book evaluation system in reflecting the quality of book content, book review indicators are adopted into the evaluation system, and different disciplines of book are applied to analyze the adaptability of the evaluation system to explore the evaluation of Chinese book quality. [**Method/process**] Sixteen indicators were selected from traditional measurement indicators, Altmetrics indicators and book review indicators to construct a comprehensive evaluation system for Chinese books empowered by the CRITIC method, while economics, computer science and physics books were selected for an empirical study to analyze the utility of book review indicators in book quality evaluation and to test the adaptability of the evaluation system in evaluating books of different disciplines. [**Result/conclusion**] The empirical results show that the comprehensive evaluation system of Chinese books has good adaptability for books of different disciplines, and the evaluation results can complement the influence and quality of books reflected by total citations and libcitations; the book review indicators can reflect the reading tendency of readers in different disciplines, and help screen out disciplinary books with high readability, scientificity and knowledge, thus providing reference for book quality evaluation activities such as reading promotion and bibliographic recommendation.

Keywords: book evaluation　quality evaluation　book review　comprehensive evaluation system　adaptability